

Harshit Awasthi

9532286965 - awasthiharshit77@gmail.com - www.linkedin.com/in/harshit-awasthi/

EMPLOYMENT HISTORY

TREDENCE - Data Engineer APR 2021 - Present

- Built big data pipelines with different data sources (i.e. file, databases and API) to handle the metadata framework in Python to build the data warehouse.
- Implemented the batch processing workflows using Apache Airflow as scheduler and Redshift as target database.
- Functionality enhancement of the existing metadata framework and adapters to accommodate the specific data requirements for solving complex big data problems.
- Designed and developed a reconciliation framework using Python to catch any anonymized/untold changes at source to maintain the data integrity between source and data lake.
- Implemented task-based trigger using Airflow sensor where a task waits on another task (in a different DAG) to complete execution.

NEENOPAL - Data Engineer SEP 2020 - APR 2021

- Worked on Finance dataset to assess business requirements and perform source-to-target data mapping. Created the backend tables in SQL Server feeding the data to Tableau BI dashboards.
- Developed scripts to perform ETL using Python. Created the end to end ETL Logic for data pipeline.
- Implemented the ETL using AWS Lambda function and S3. Created the data pipeline ingesting data from S3 to MySQL database with the business logic for an e-Commerce client. Scheduling the run using AWS Cloud Watch.

QUAERO - Associate Data Engineer JAN 2019 - MAY 2020

- Worked on Quaero CDP (A PaaS product for managing first-party data) for creating Aggregation and Transformation workflows for transforming incoming data from multiple sources.
- Worked on ESPN AdScope and Fan Relationship Management project to create data pipeline ingesting data from different data sources such as website click-stream, Google Double click Ad
- Impressions and Mobile App Subscriptions then processing data using workflows for client-facing deliverable data on Snowflake.
- Built queries to ensure data quality & data cleaning and identifying data gaps & inconsistencies.
- Developed automated reports for daily load monitoring using SQL Server and designed client-facing reports and dashboards using Looker software.

VEDANTA RESOURCES - Electrical Engineer JUN 2017 - DEC 2018

- Strategic planning for electrical installations so as to support underground mines development heavy machinery power supply units, ventilation and de-watering electrical units.
- Monitored and analysed the data of process parameters of roaster plant in order to increase the performance of wet gas precipitator and reduce the downtime using Excel.

SKILLS

Languages : SQL, Python

Databases : SQL Server, Snowflake, Redshift, Hive

Libraries : Numpy, Pandas, AWS Wrangler, SQL Alchemy

Other Technologies : Airflow, AWS Cloud (Lambda, S3)

Version Control : Bit Bucket

EDUCATION

- **B.TECH** Electrical Engineering from NIT Agartala (7.61 CGPA) | 2013-2017
- **12th** BNSD Shiksha Niketan Inter College, UP Board with 84.2% | 2012
- **10th** BNSD Shiksha Niketan Inter College, UP Board with 82.83% | 2010